FIG. 1

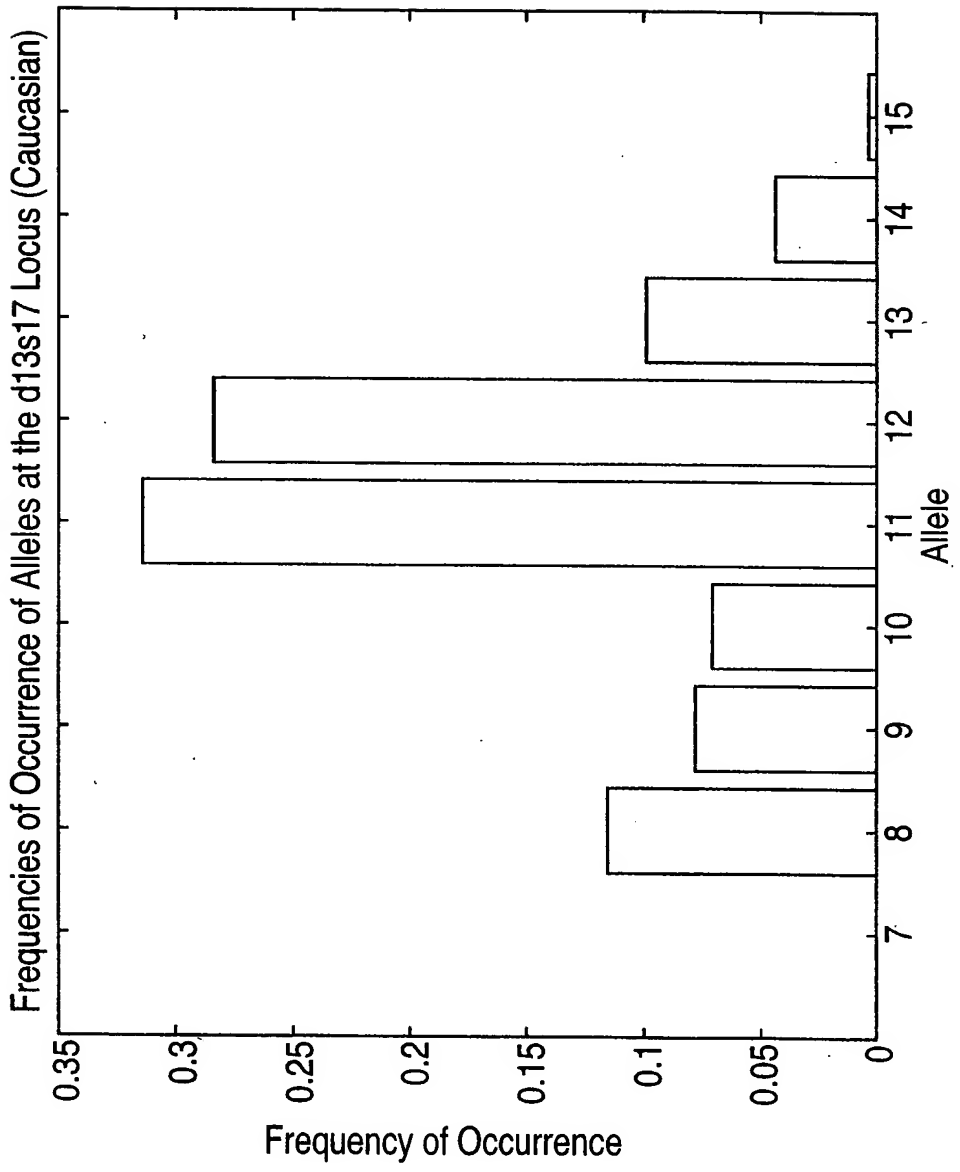

Frequencies of Occurrence of Alleles at the d13s17 Locus (Caucasian)
Frequency of Occurrence
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
7
8
9
10
11
12
13
14
15
Allele

FIG. 2

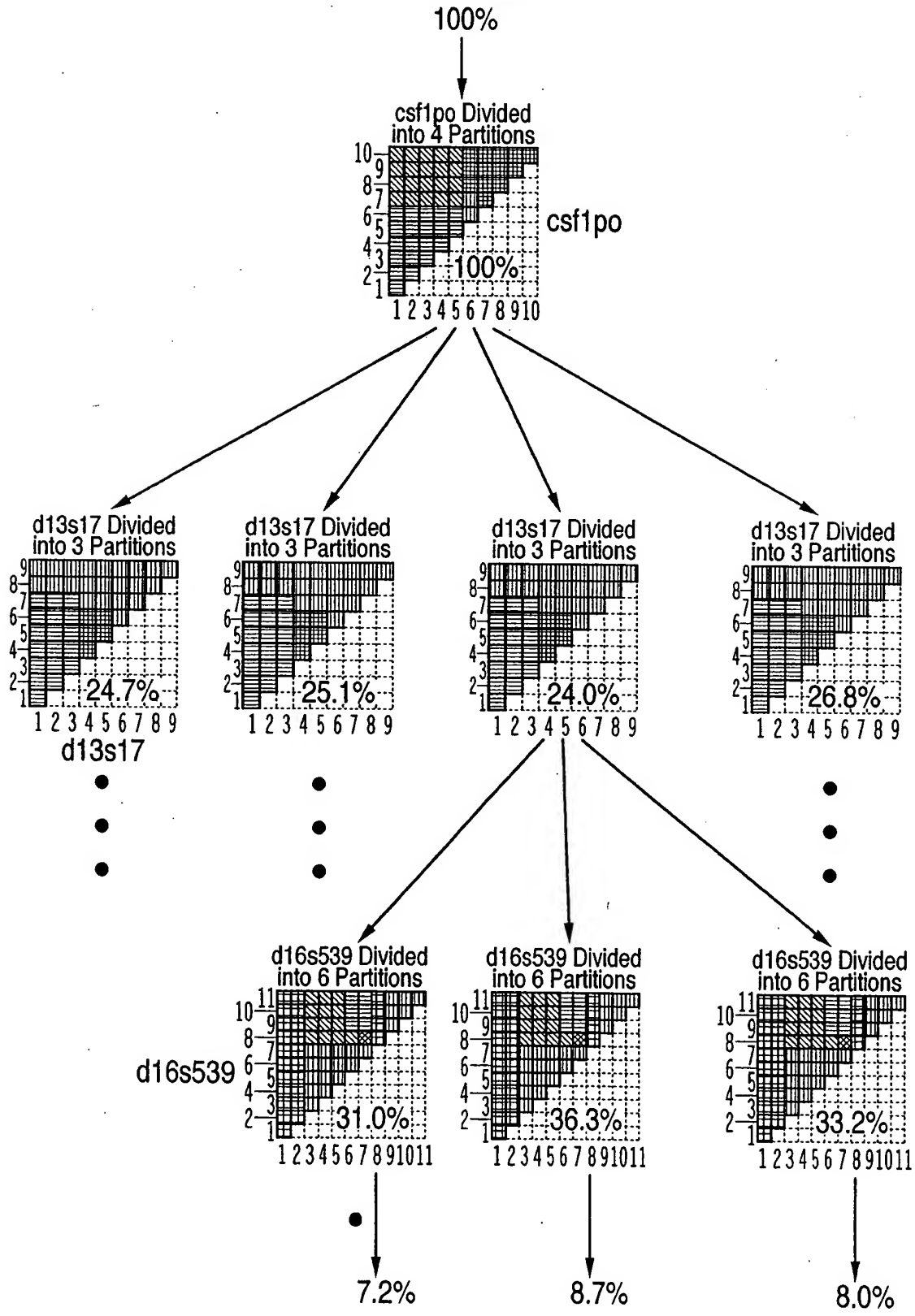

100%
csf1po Divided into 4 Partitions
100%
csf1po
d13s17 Divided into 3 Partitions
24.7%
d13s17
d13s17 Divided into 3 Partitions
25.1%
d13s17 Divided into 3 Partitions
24.0%
d13s17 Divided into 3 Partitions
26.8%
d16s539 Divided into 6 Partitions
31.0%
d16s539
d16s539 Divided into 6 Partitions
36.3%
d16s539 Divided into 6 Partitions
33.2%
7.2%
8.7%
8.0%

FIG. 3

SINGLE HOST SEARCH SERVER

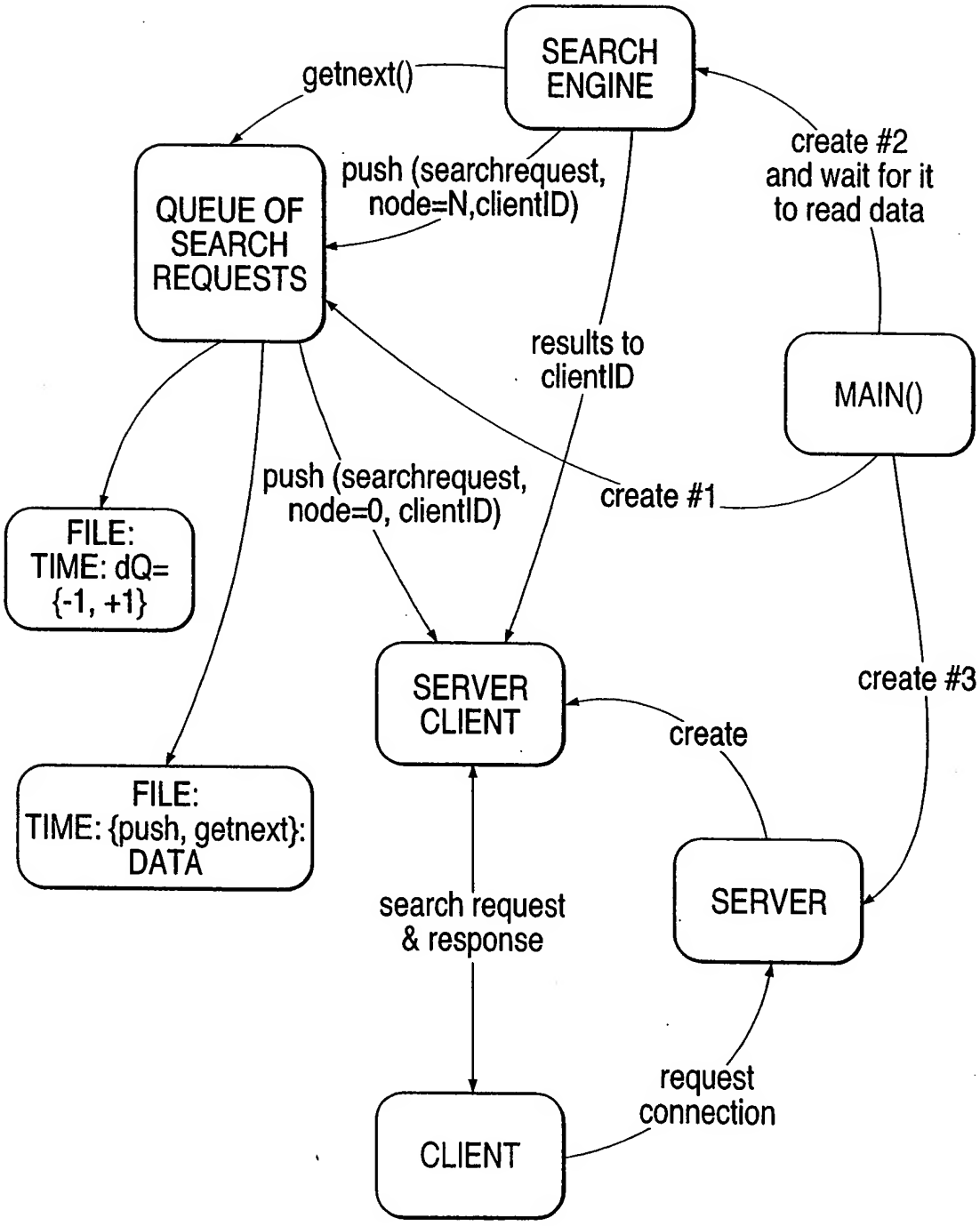


FIG. 4

d16s539 Divided into 6 Partitions

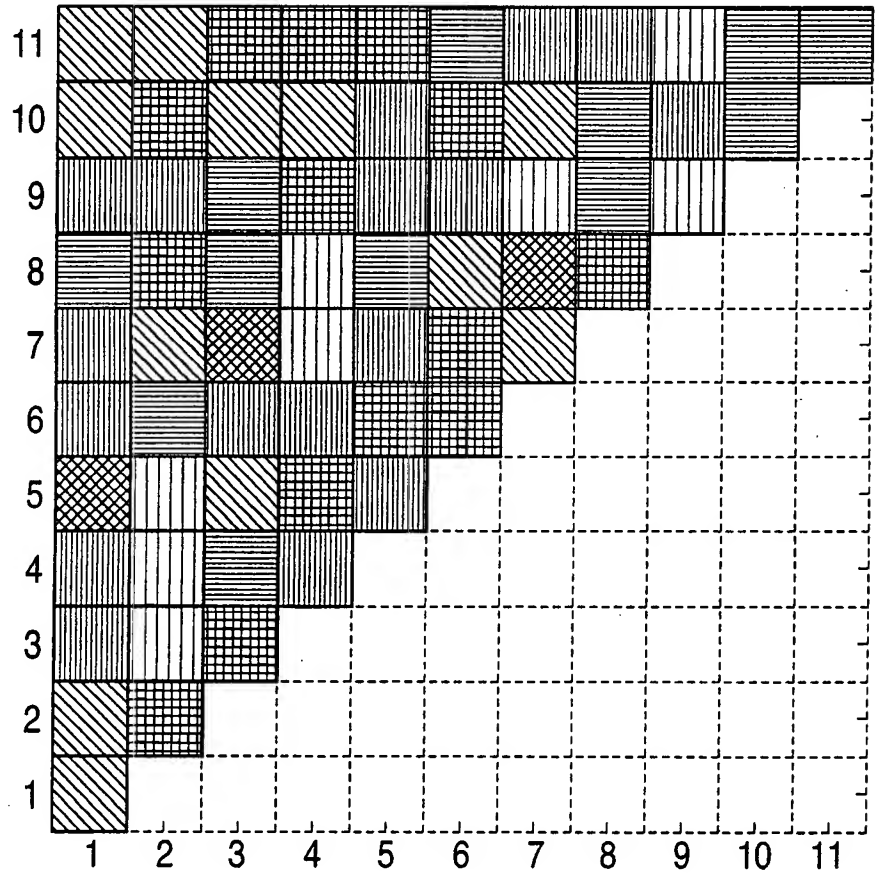


FIG. 5

d16s539 Divided into 6 Partitions

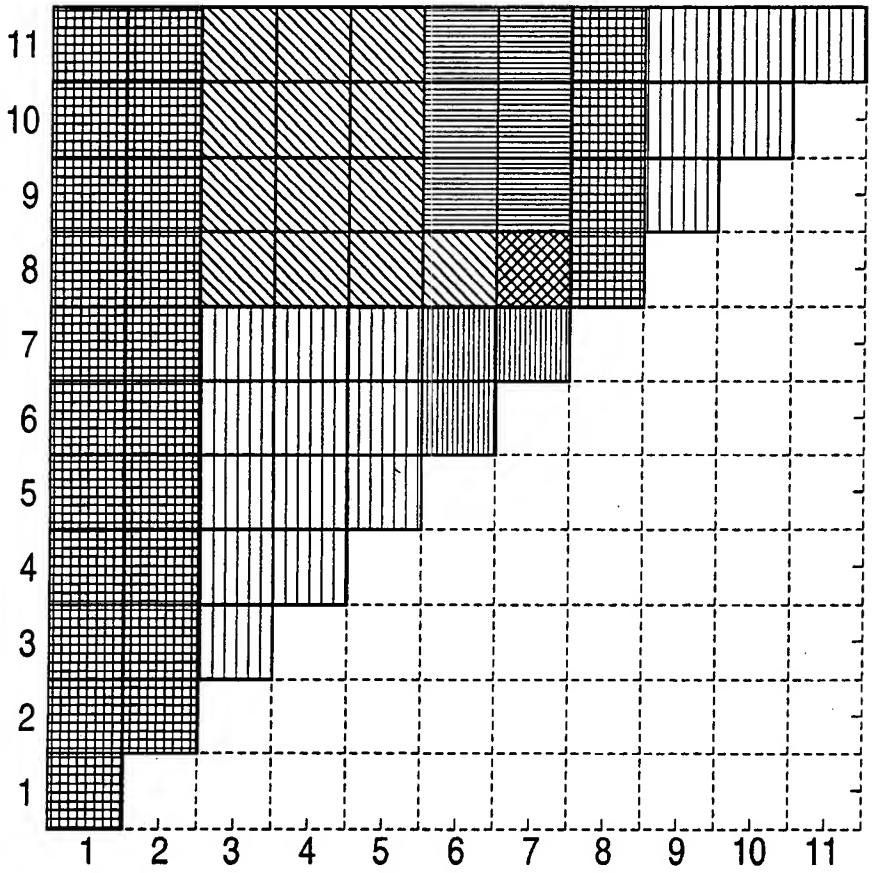


FIG. 6

d16s539 Divided into 5 Partitions

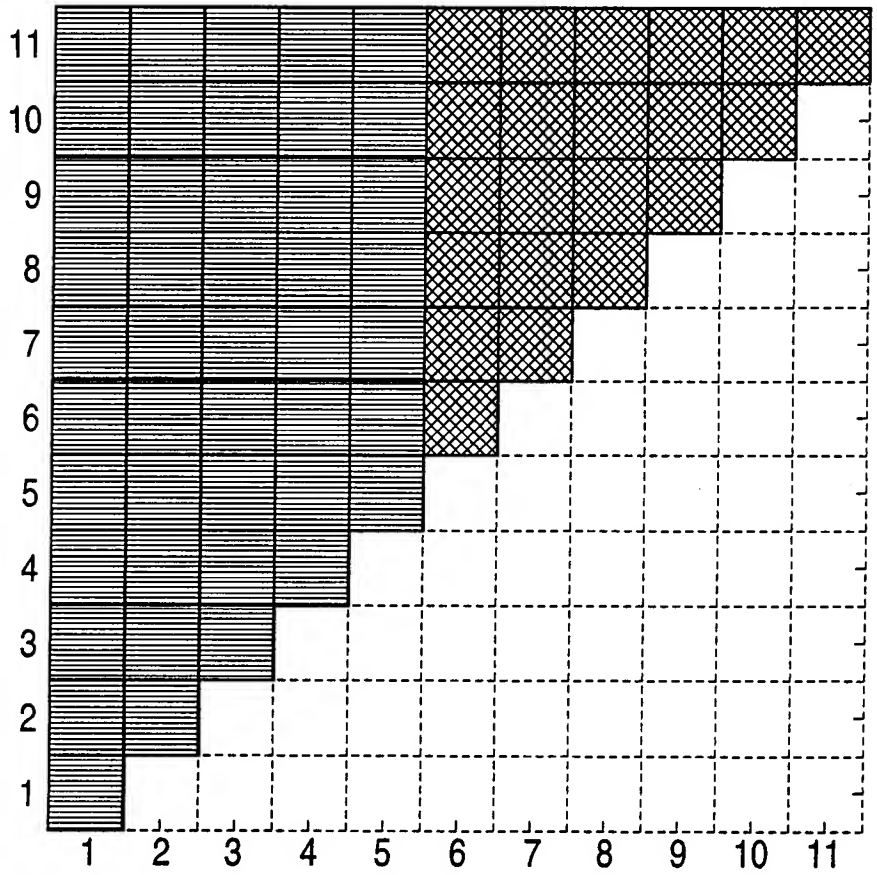


FIG. 7

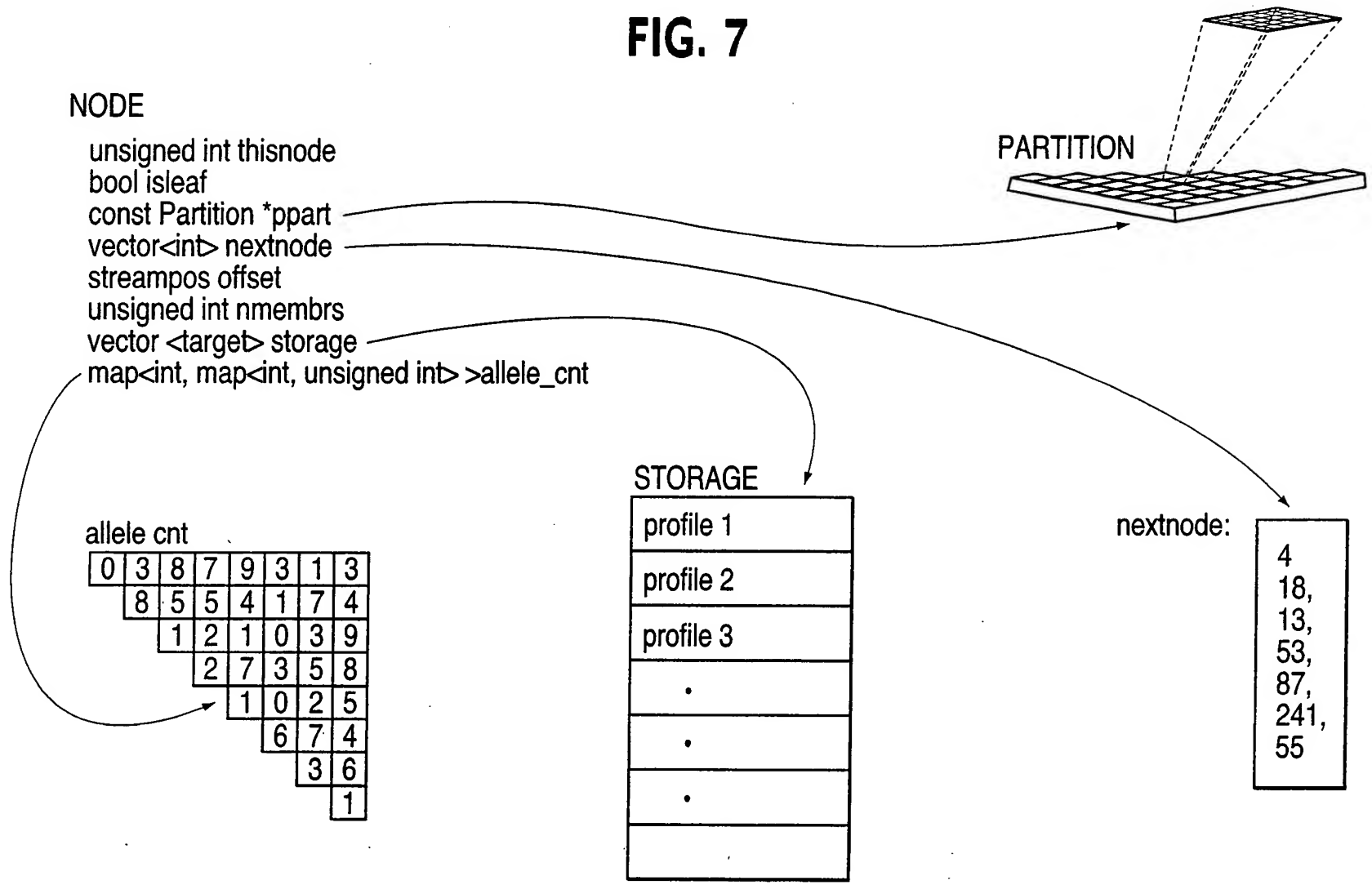

NODE
unsigned int thisnode
bool isleaf
const Partition *ppart
vector<int> nextnode
streampos offset
unsigned int nmembrs
vector <target> storage
map<int, map<int, unsigned int> >allele_cnt
PARTITION
allele cnt
0 3 8 7 9 3 1 3
8 5 5 4 1 7 4
1 2 1 0 3 9
2 7 3 5 8
1 0 2 5
6 7 4
3 6
1
STORAGE
profile 1
profile 2
profile 3
nextnode:
4
18,
13,
53,
87,
241,
55

FIG. 8

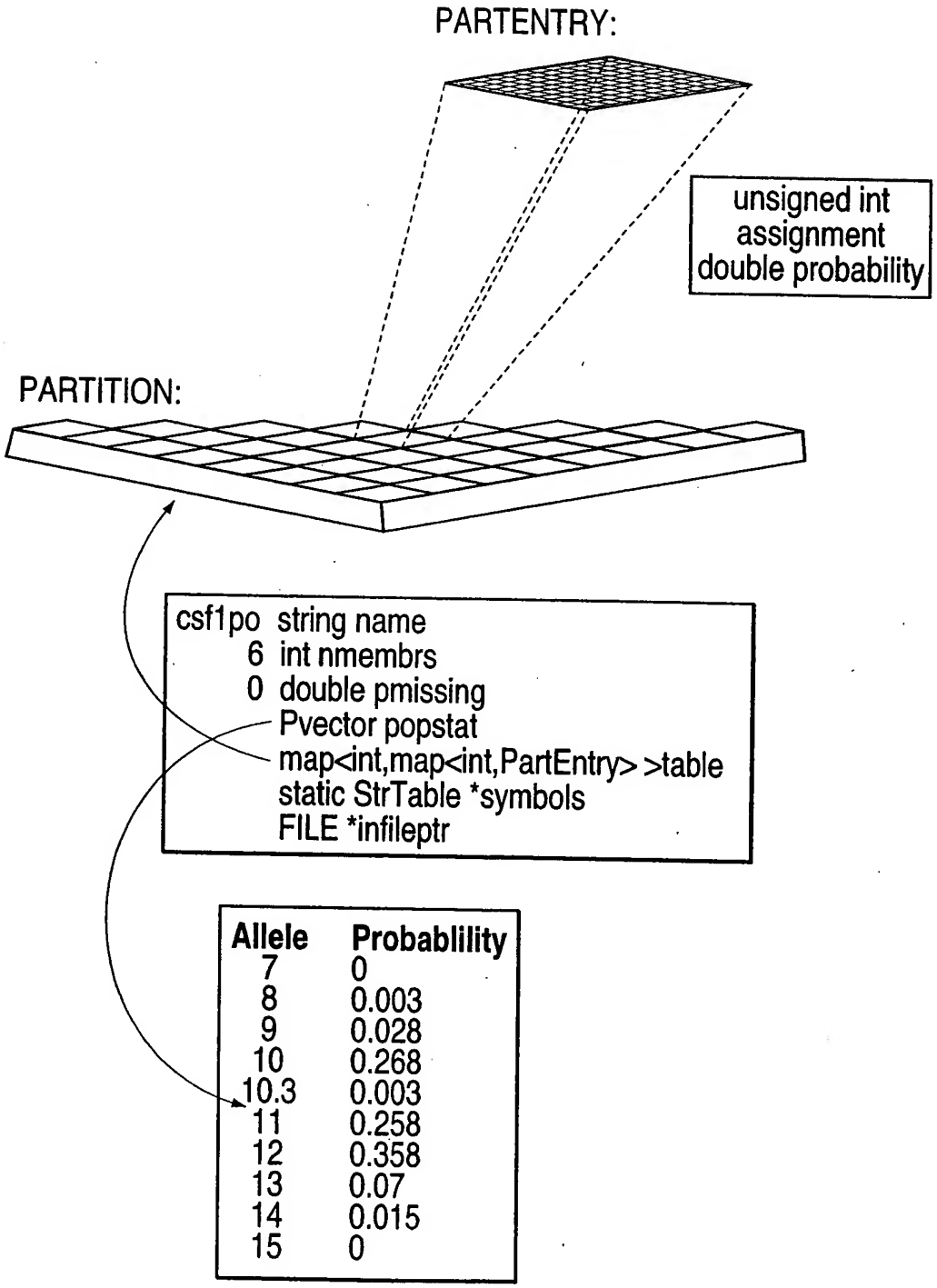

PARTENTRY:
unsigned int
assignment
double probability
PARTITION:
csf1po string name
6 int nmembrs
0 double pmissing
Pvector popstat
map<int,map<int,PartEntry> >table
static StrTable *symbols
FILE *infileptr
Allele Probablility
7 0
8 0.003
9 0.028
10 0.268
10.3 0.003
11 0.258
12 0.358
13 0.07
14 0.015
15 0

FIG. 9

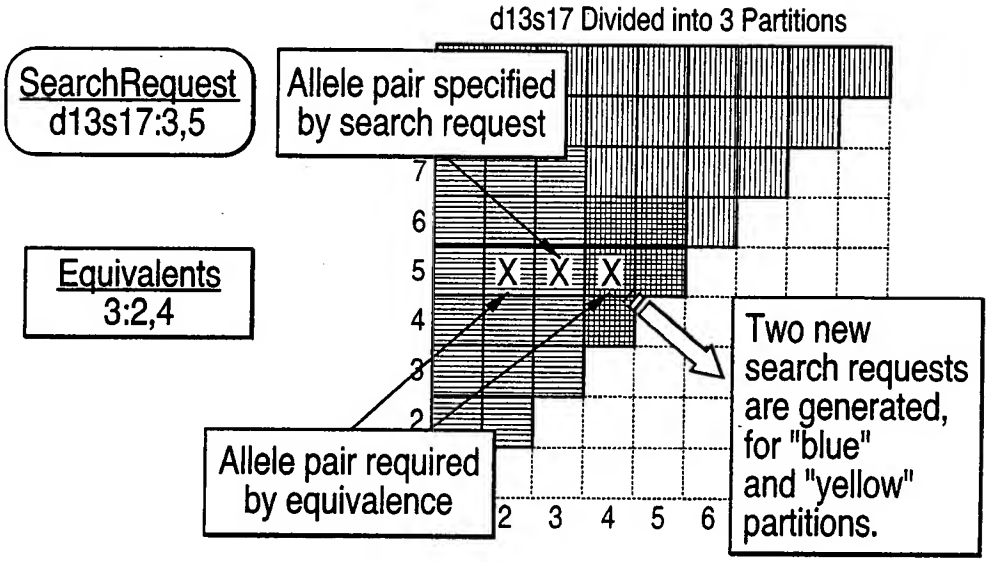

d13s17 Divided into 3 Partitions
SearchRequest
d13s17:3,5
Allele pair specified by search request
Equivalents
3:2,4
Allele pair required by equivalence
Two new search requests are generated, for "blue" and "yellow" partitions.
7
6
5
4
3
2
X
X
X
2
3
4
5
6

FIG. 10

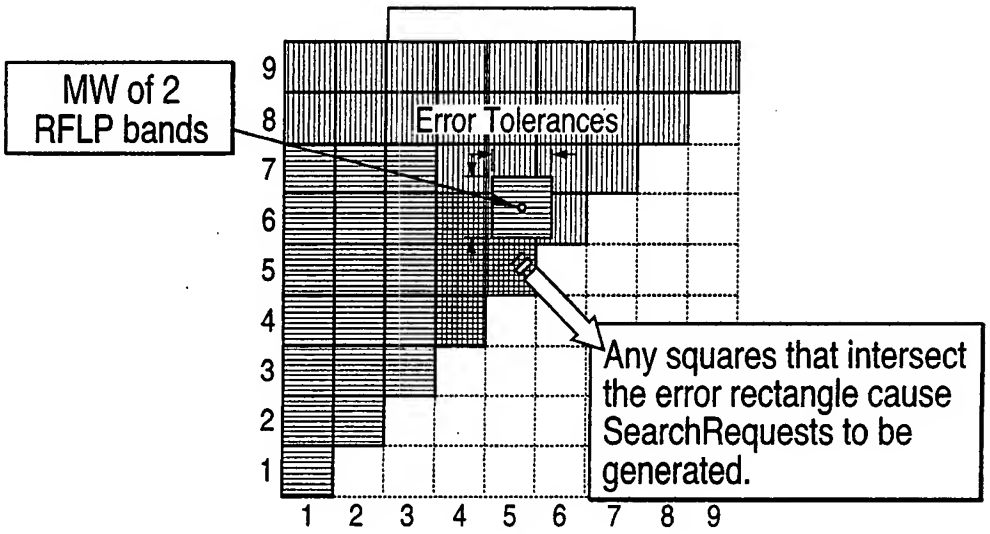

MW of 2
RFLP bands
Error Tolerances
Any squares that intersect the error rectangle cause SearchRequests to be generated.
9
8
7
6
5
4
3
2
1
1 2 3 4 5 6 7 8 9

FIG. 11

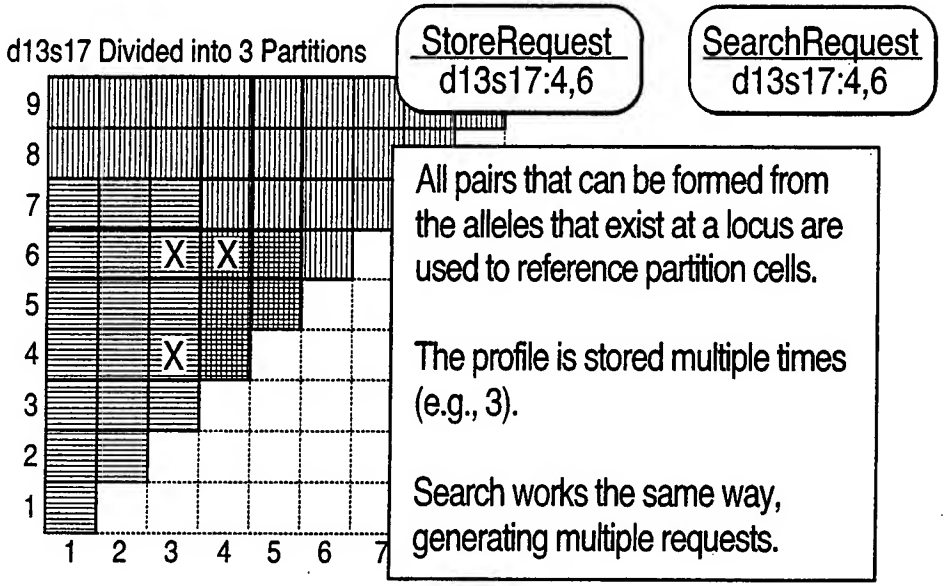

d13s17 Divided into 3 Partitions
StoreRequest
d13s17:4,6
SearchRequest
d13s17:4,6
9
8
7
6
5
4
3
2
1
X
X
X
1 2 3 4 5 6 7
All pairs that can be formed from the alleles that exist at a locus are used to reference partition cells.
The profile is stored multiple times (e.g., 3).
Search works the same way, generating multiple requests.

FIG. 12

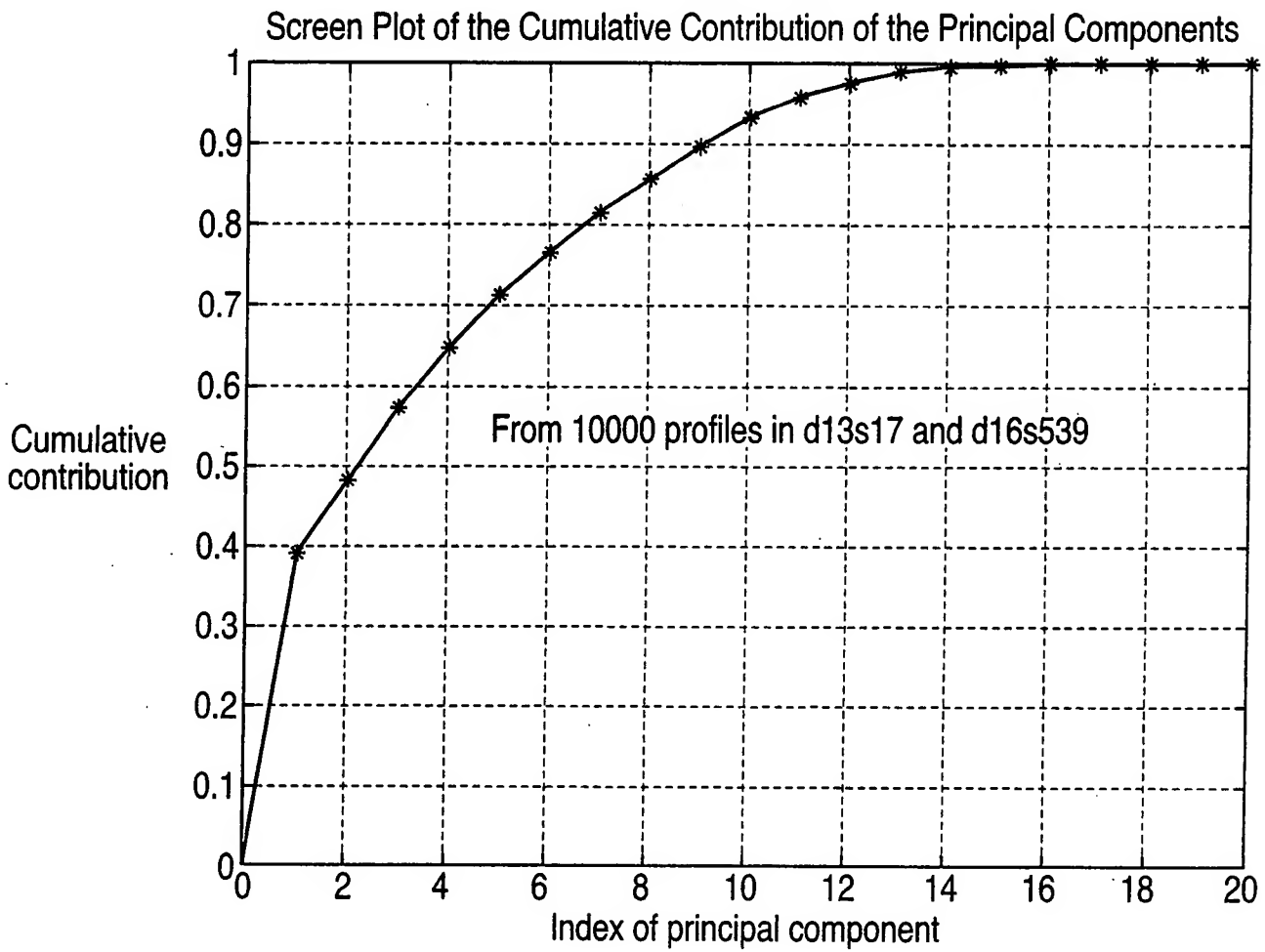

Screen Plot of the Cumulative Contribution of the Principal Components
Cumulative contribution
From 10000 profiles in d13s17 and d16s539
Index of principal component
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
2
4
6
8
10
12
14
16
18
20

FIG. 13

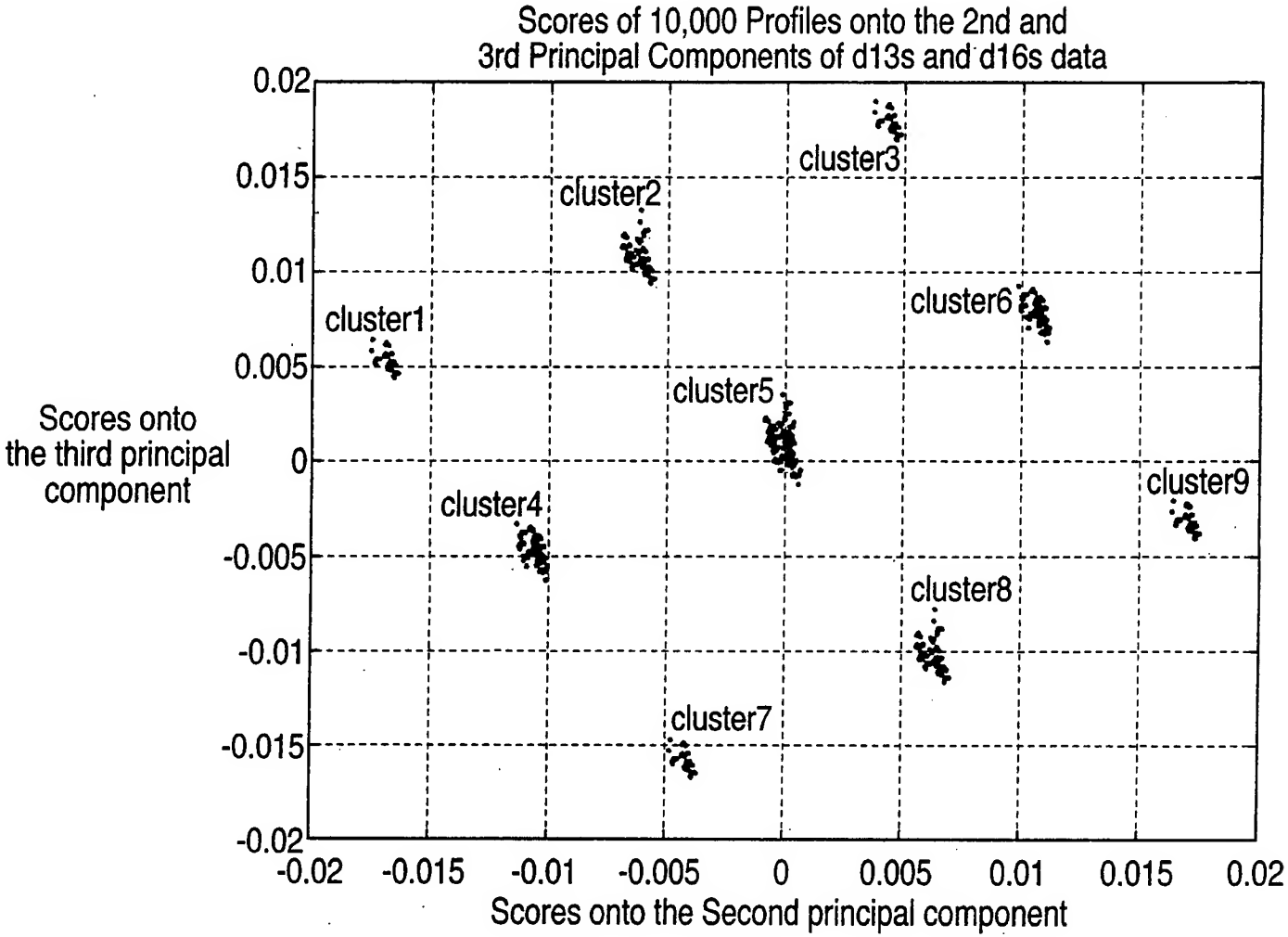

Scores of 10,000 Profiles onto the 2nd and
3rd Principal Components of d13s and d16s data
Scores onto
the third principal
component
0.02
0.015
0.01
0.005
0
-0.005
-0.01
-0.015
-0.02
cluster1
cluster2
cluster3
cluster4
cluster5
cluster6
cluster7
cluster8
cluster9
-0.02
-0.015
-0.01
-0.005
0
0.005
0.01
0.015
0.02
Scores onto the Second principal component

FIG. 14

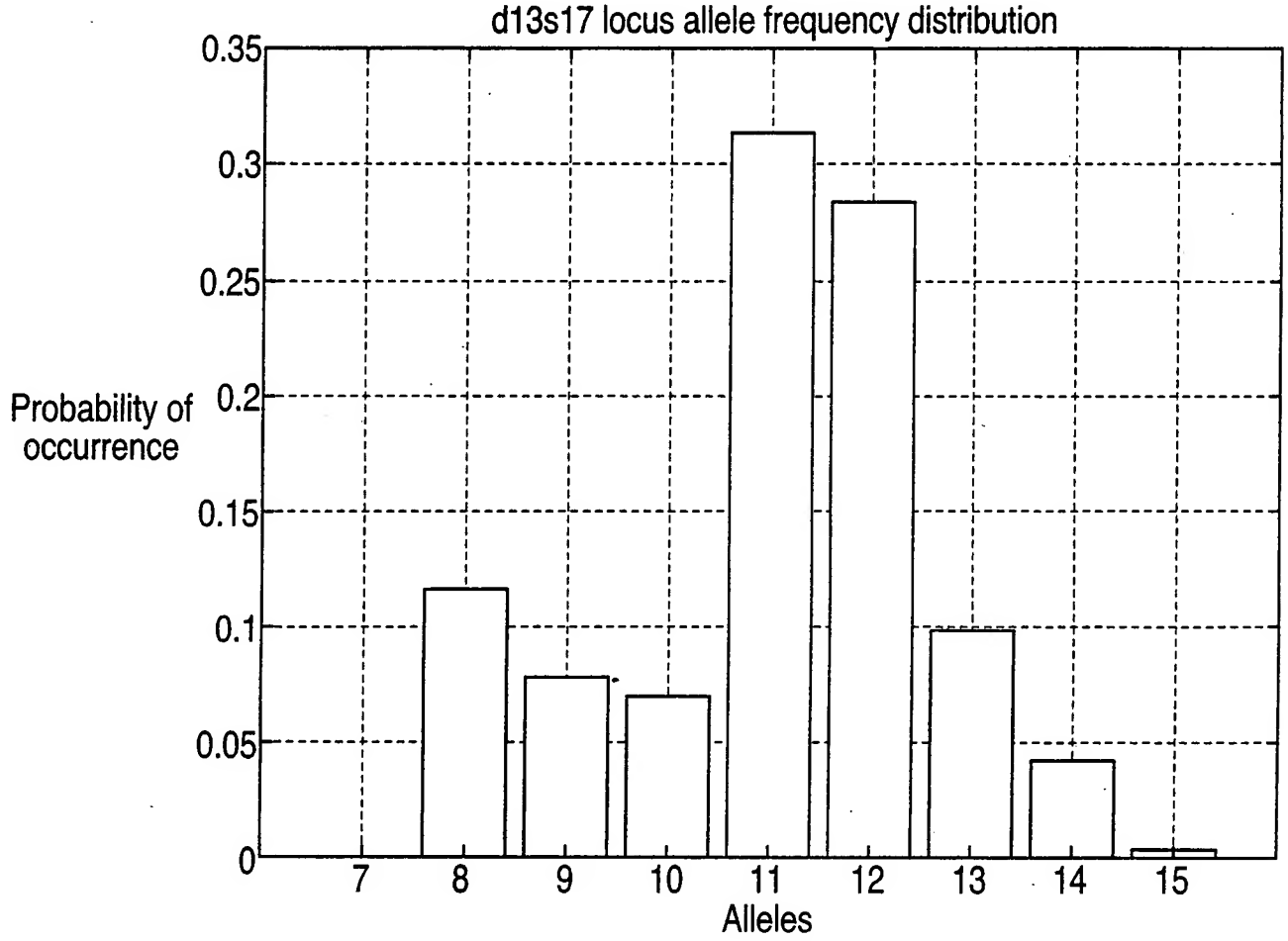

d13s17 locus allele frequency distribution
Probability of occurrence
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
7
8
9
10
11
12
13
14
15
Alleles

FIG. 15

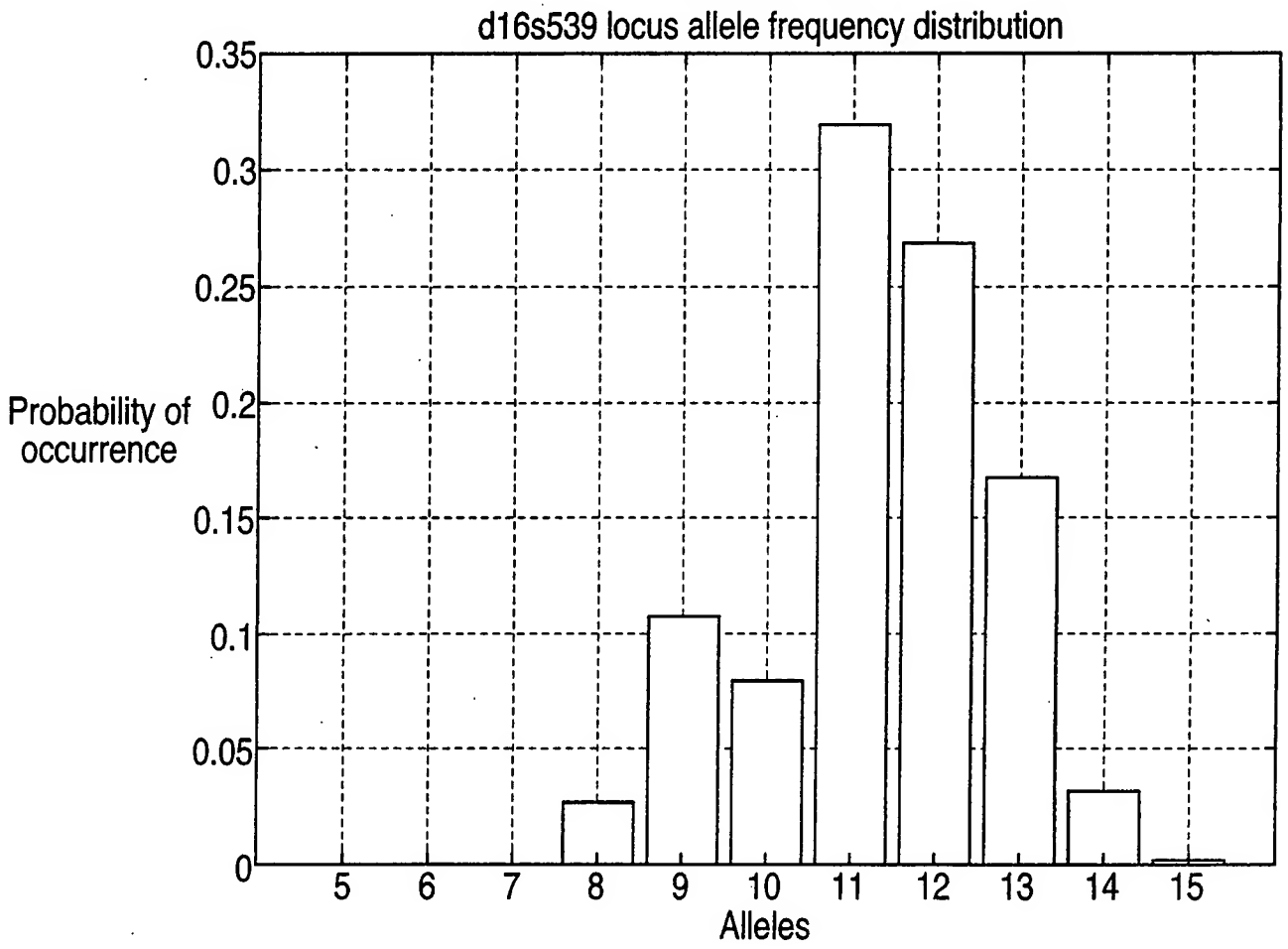

d16s539 locus allele frequency distribution
Probability of occurrence
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
5
6
7
8
9
10
11
12
13
14
15
Alleles

FIG. 16

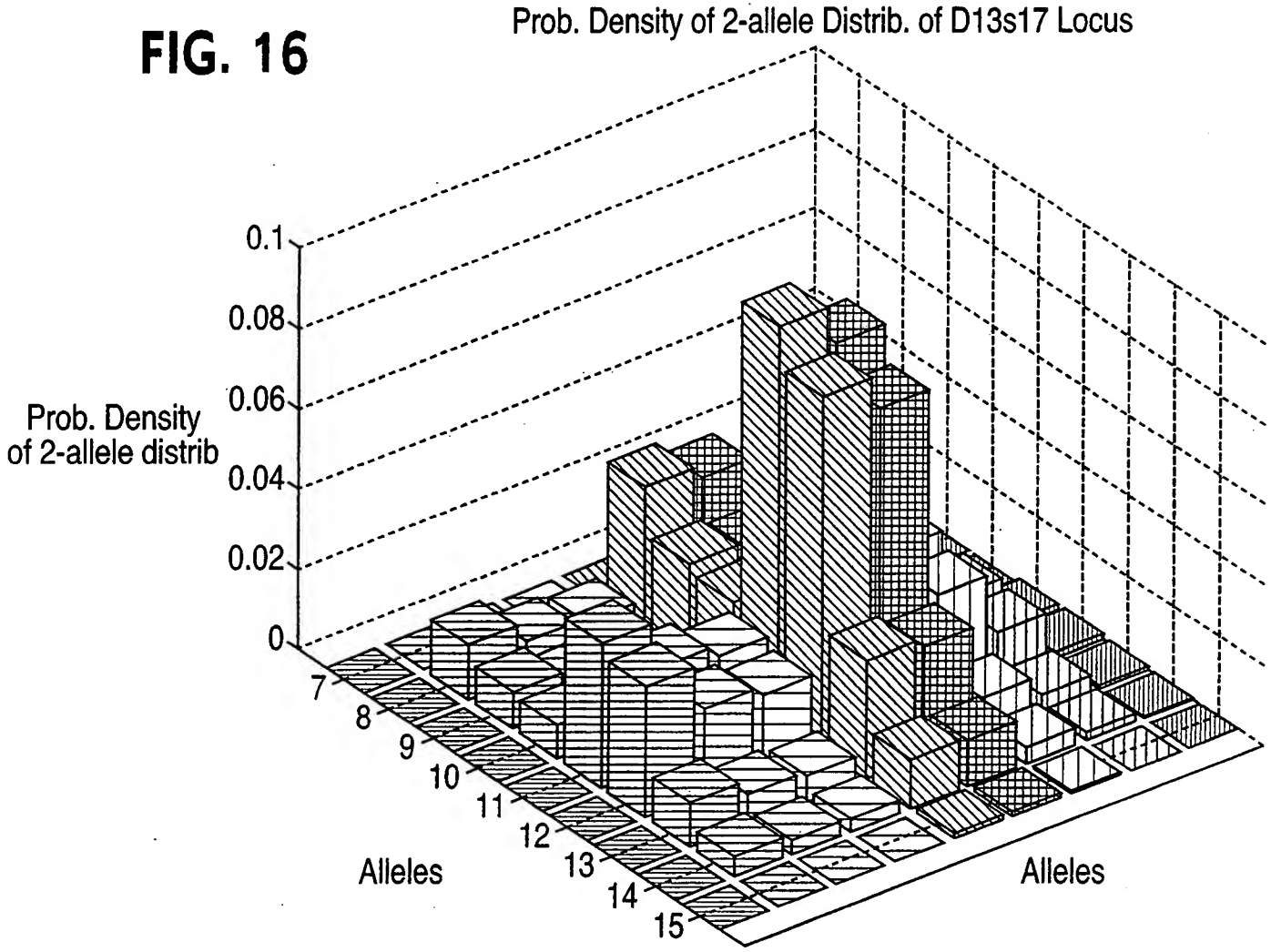

Prob. Density of 2-allele Distrib. of D13s17 Locus
Prob. Density of 2-allele distrib
0.1
0.08
0.06
0.04
0.02
0
7
8
9
10
11
12
13
14
15
Alleles
Alleles

FIG. 17

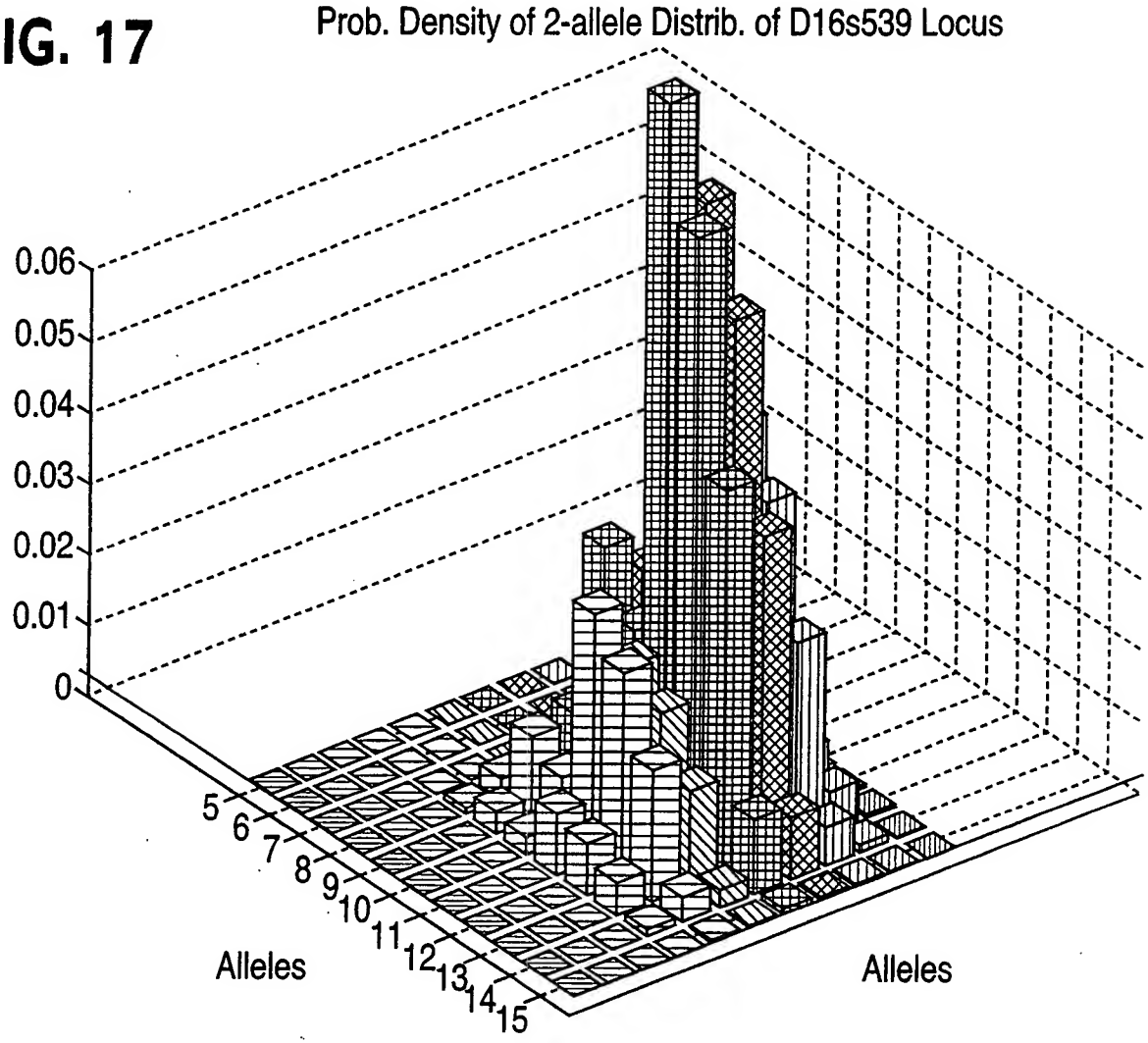

Prob. Density of 2-allele Distrib. of D16s539 Locus
0.06
0.05
0.04
0.03
0.02
0.01
0
5
6
7
8
9
10
11
12
13
14
15
Alleles
Alleles

FIG. 18

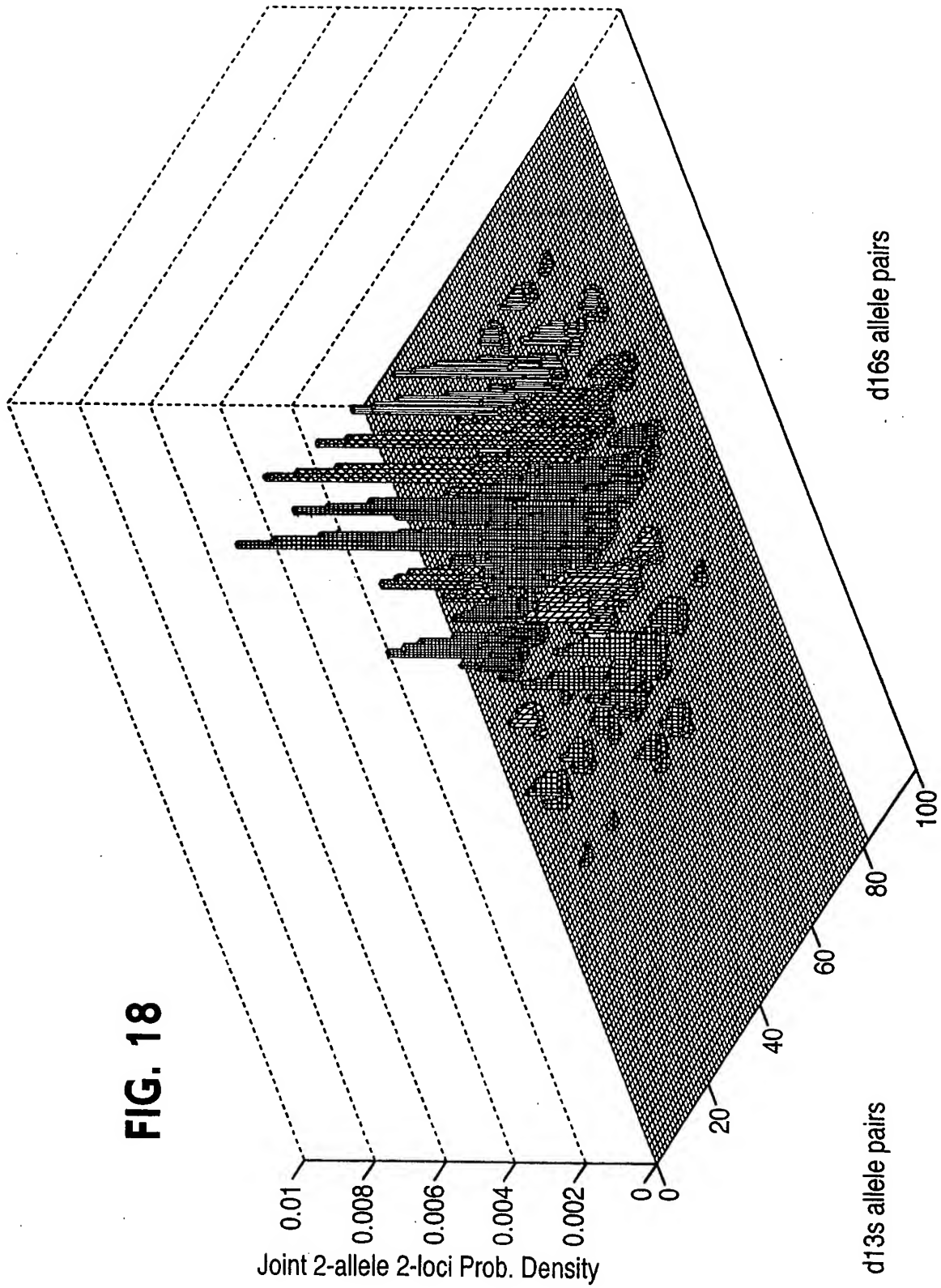

Joint 2-allele 2-loci Prob. Density
0.01
0.008
0.006
0.004
0.002
0
0
20
40
60
80
100
d13s allele pairs
d16s allele pairs

FIG. 19

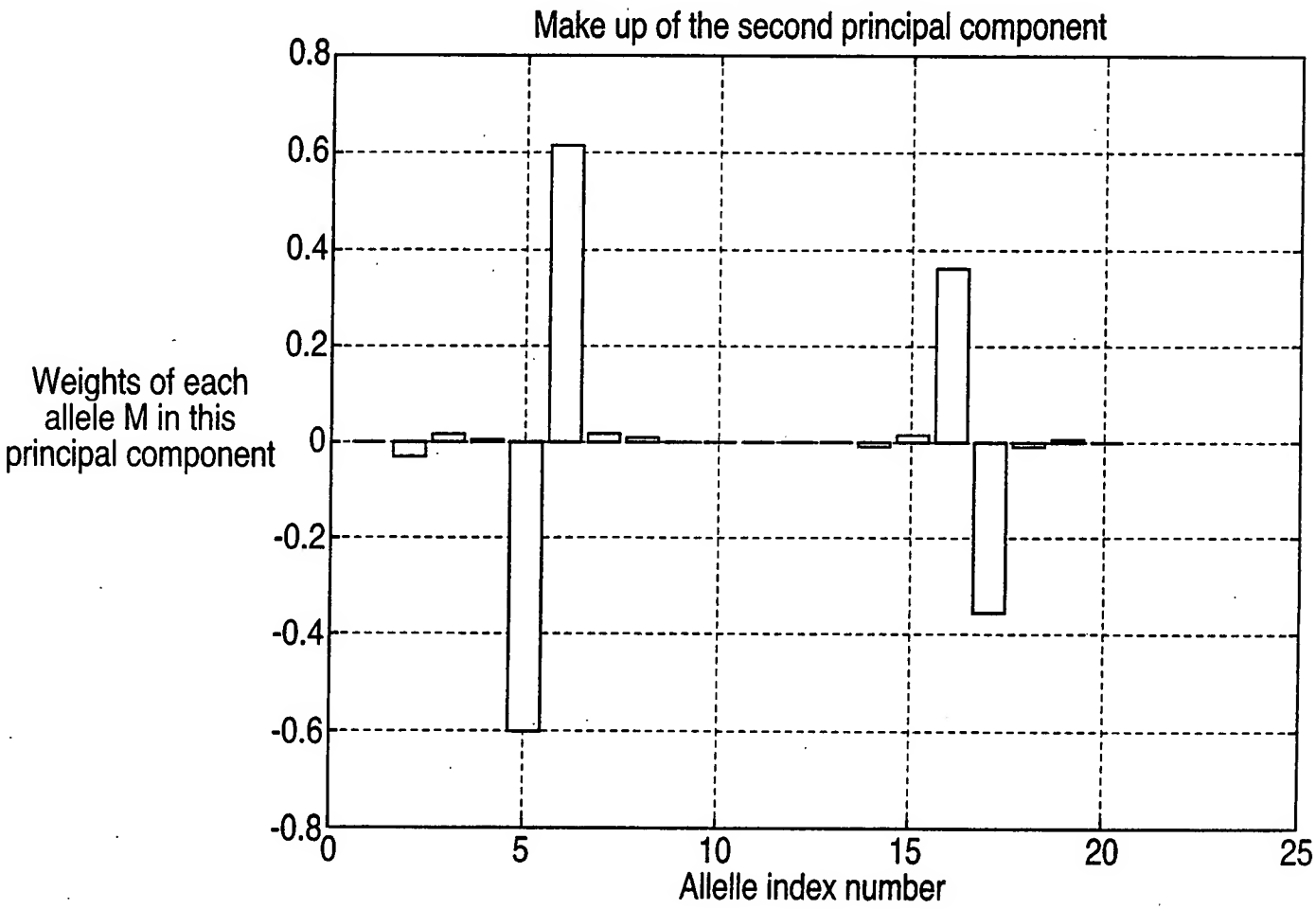

Make up of the second principal component
Weights of each allele M in this principal component
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
0
5
10
15
20
25
Allelle index number

FIG. 20

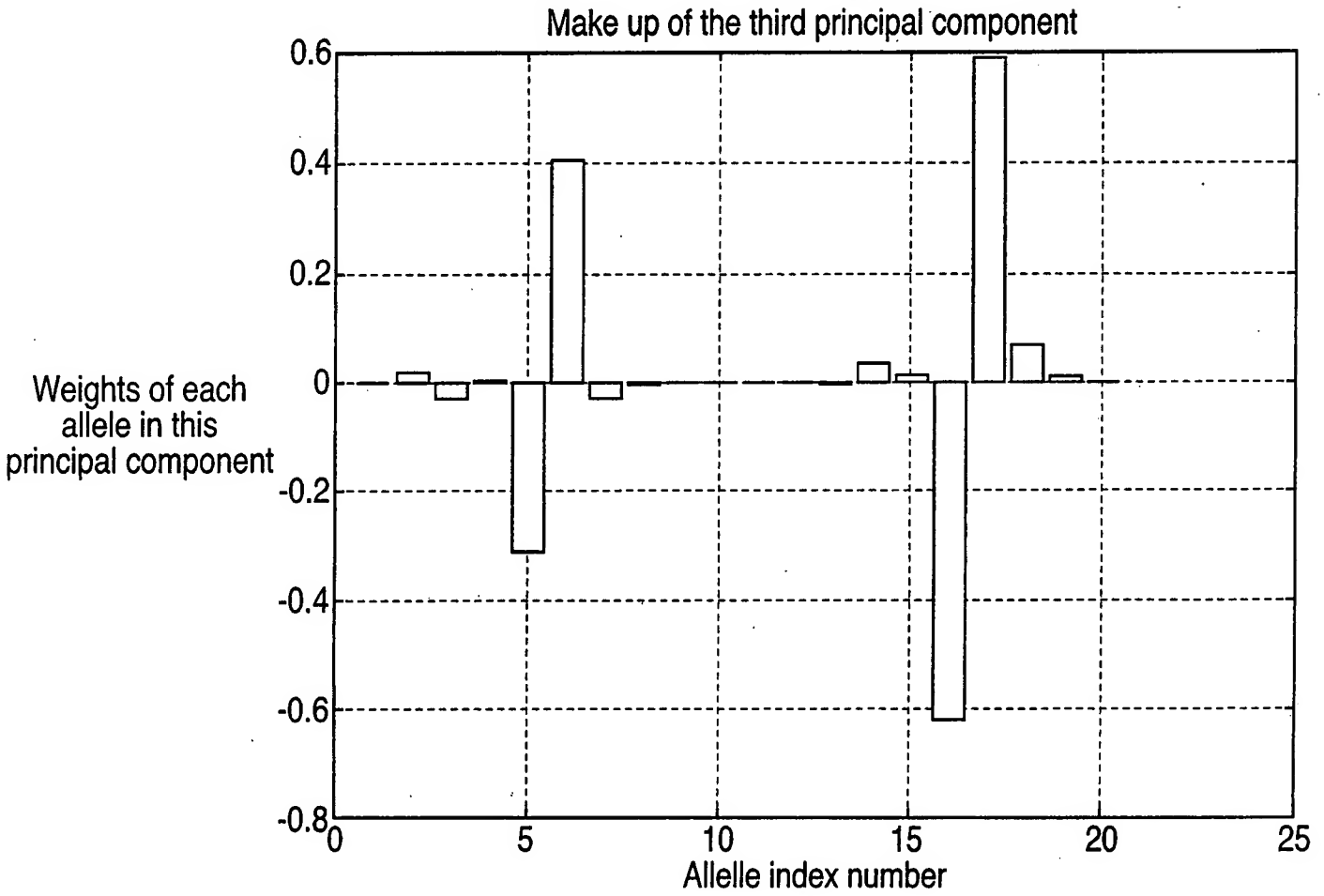

Make up of the third principal component
Weights of each allele in this principal component
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
0
5
10
15
20
25
Allelle index number

FIG. 21A

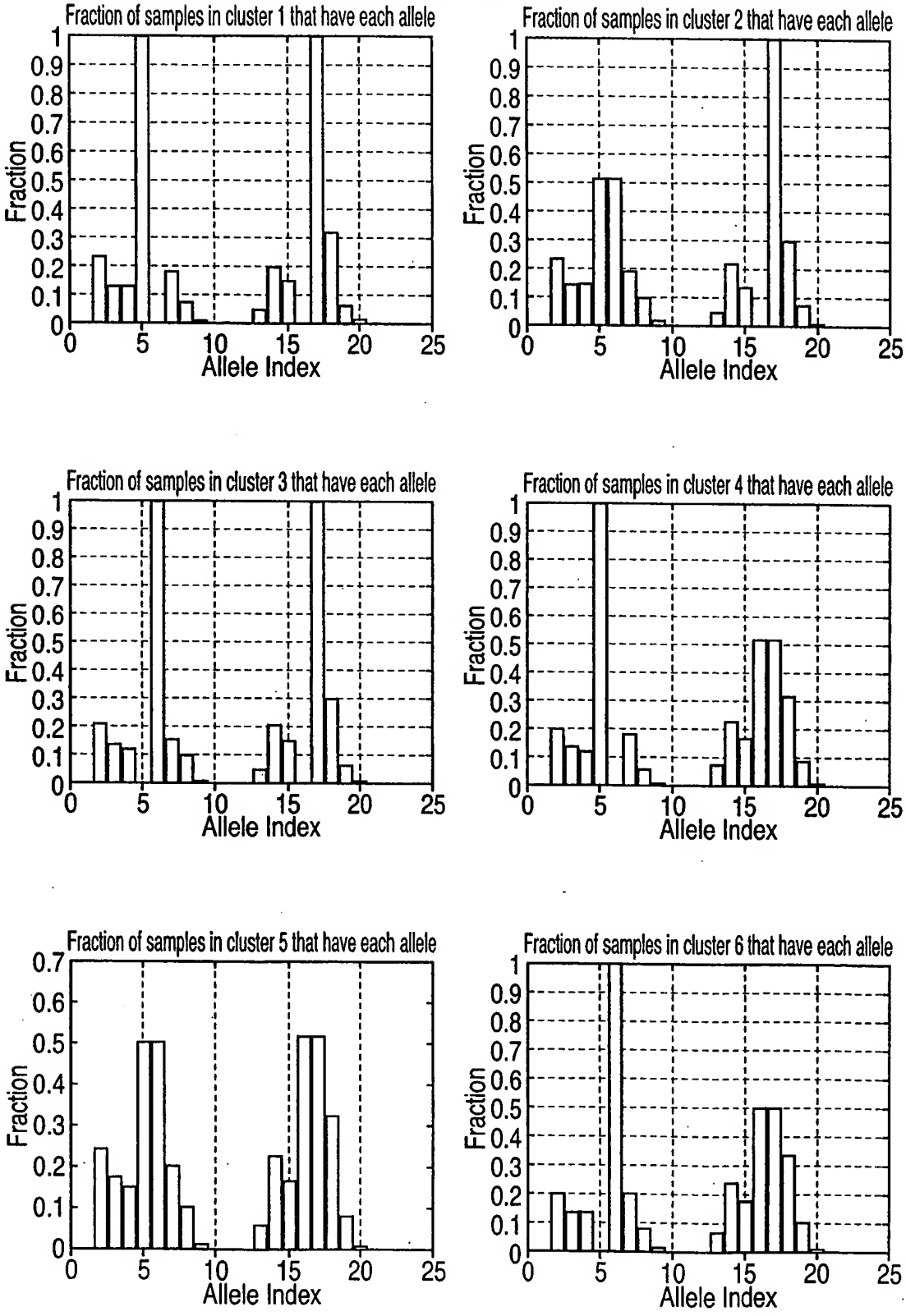
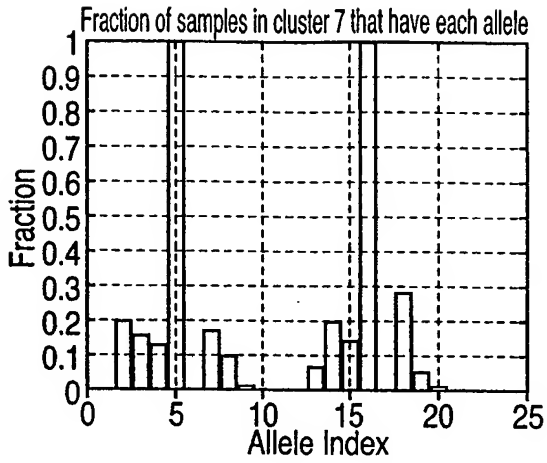
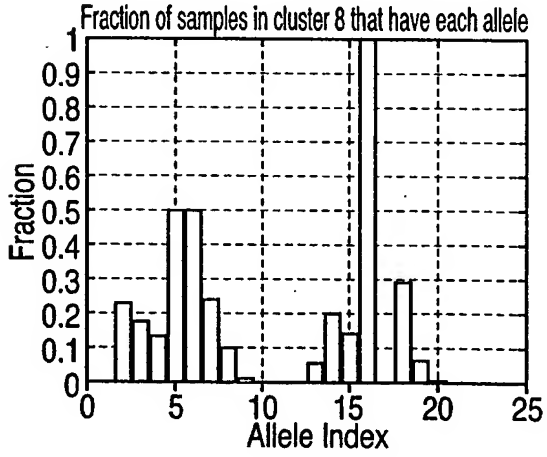

Fraction of samples in cluster 1 that have each allele
Fraction of samples in cluster 2 that have each allele
Fraction of samples in cluster 3 that have each allele
Fraction of samples in cluster 4 that have each allele
Fraction of samples in cluster 5 that have each allele
Fraction of samples in cluster 6 that have each allele
Fraction
Allele Index

FIG. 21B


Fraction of samples in cluster 7 that have each allele
Fraction
Allele Index
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
5
10
15
20
25


Fraction of samples in cluster 8 that have each allele
Fraction
Allele Index
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
5
10
15
20
25

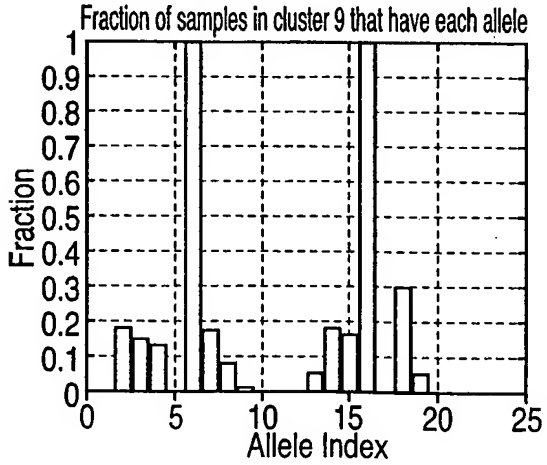

Fraction of samples in cluster 9 that have each allele
Fraction
Allele Index
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
5
10
15
20
25

FIG. 22

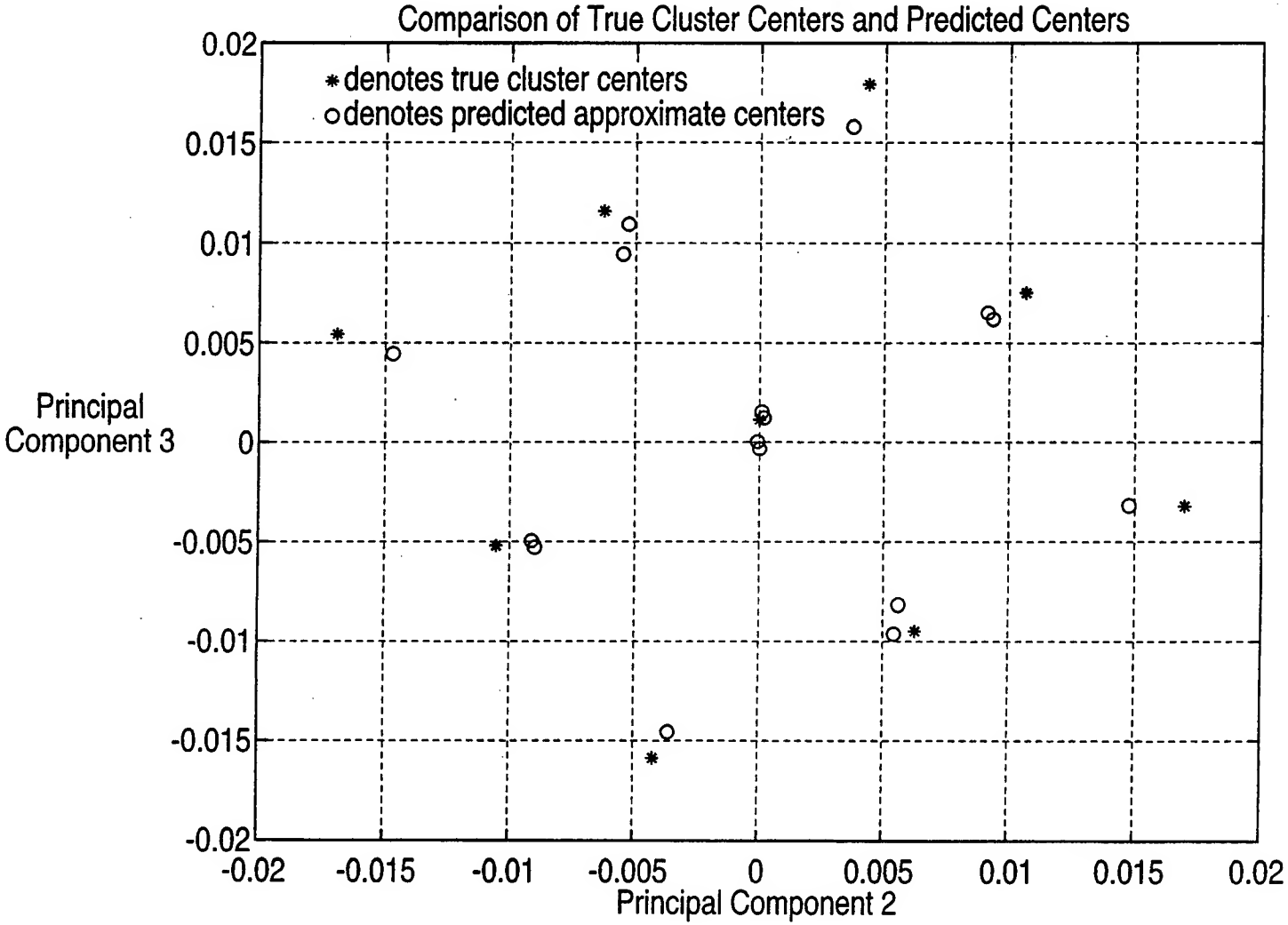

Comparison of True Cluster Centers and Predicted Centers
*denotes true cluster centers
odenotes predicted approximate centers
Principal Component 3
Principal Component 2
0.02
0.015
0.01
0.005
0
-0.005
-0.01
-0.015
-0.02
-0.02
-0.015
-0.01
-0.005
0
0.005
0.01
0.015
0.02

FIG. 23

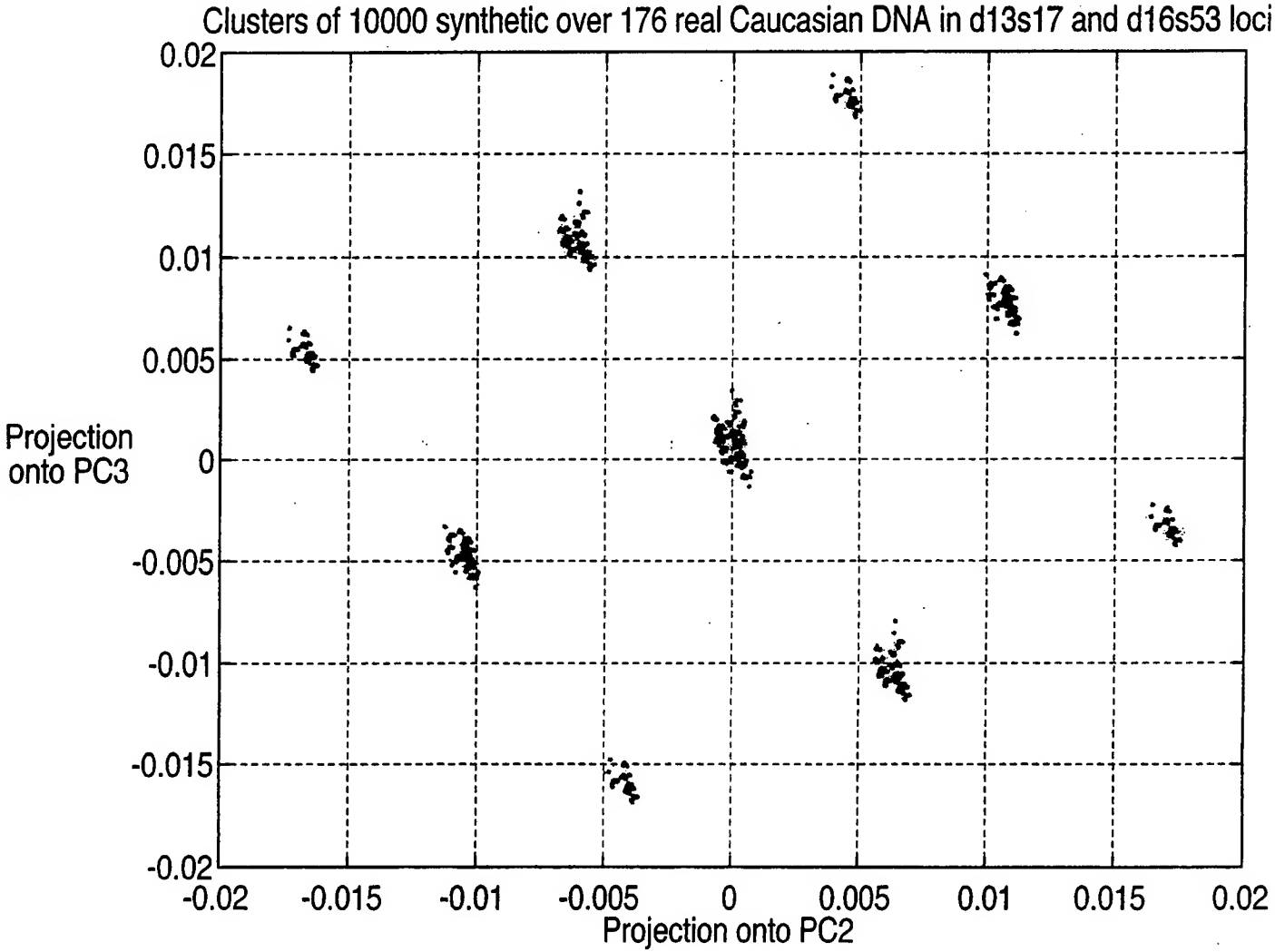

Clusters of 10000 synthetic over 176 real Caucasian DNA in d13s17 and d16s53 loci
Projection onto PC3
0.02
0.015
0.01
0.005
0
-0.005
-0.01
-0.015
-0.02
-0.02
-0.015
-0.01
-0.005
0
0.005
0.01
0.015
0.02
Projection onto PC2

FIG. 24

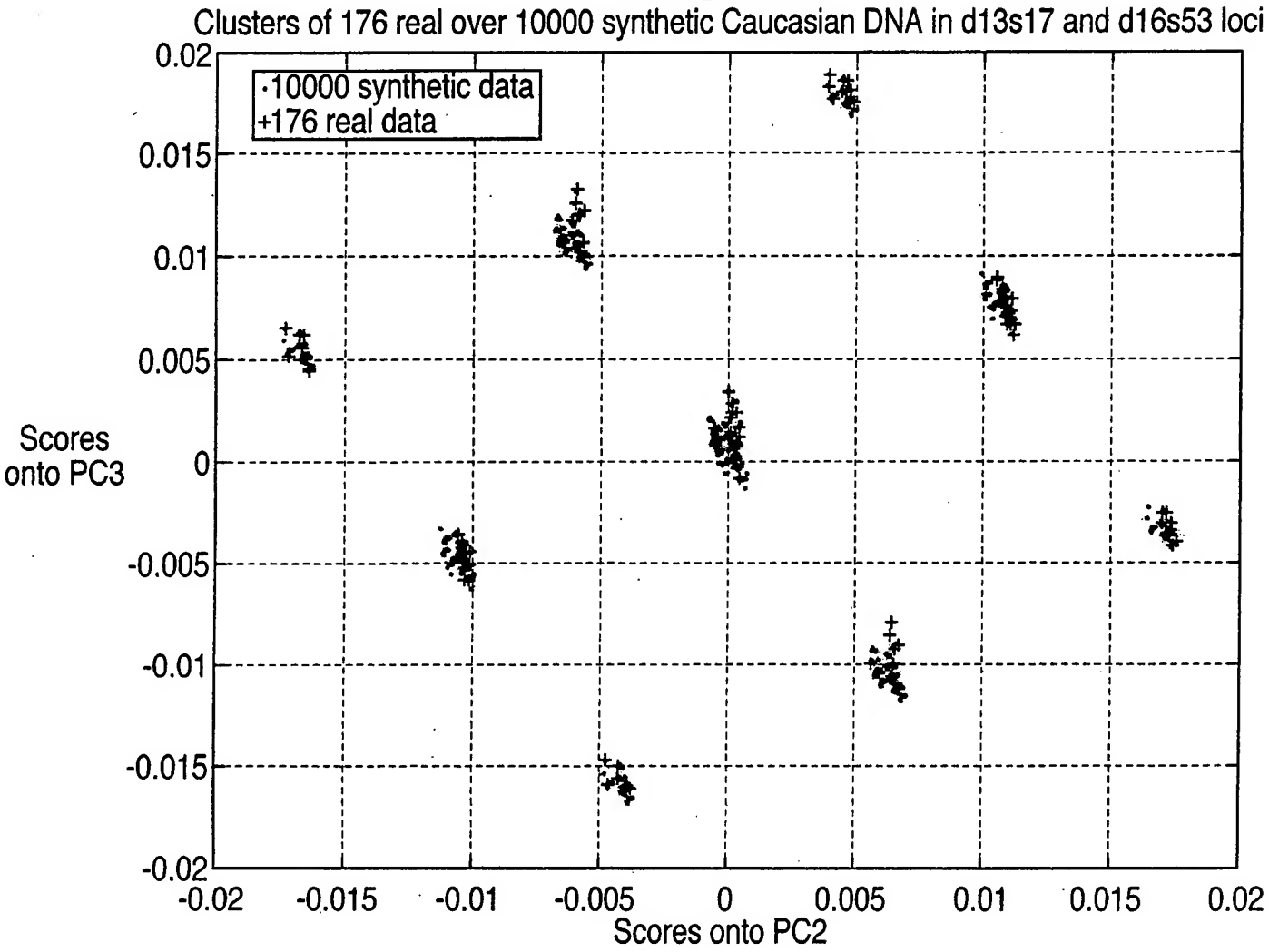

Clusters of 176 real over 10000 synthetic Caucasian DNA in d13s17 and d16s53 loci
·10000 synthetic data
+176 real data
Scores onto PC3
Scores onto PC2
0.02
0.015
0.01
0.005
0
-0.005
-0.01
-0.015
-0.02
-0.02
-0.015
-0.01
-0.005
0
0.005
0.01
0.015
0.02

FIG. 25

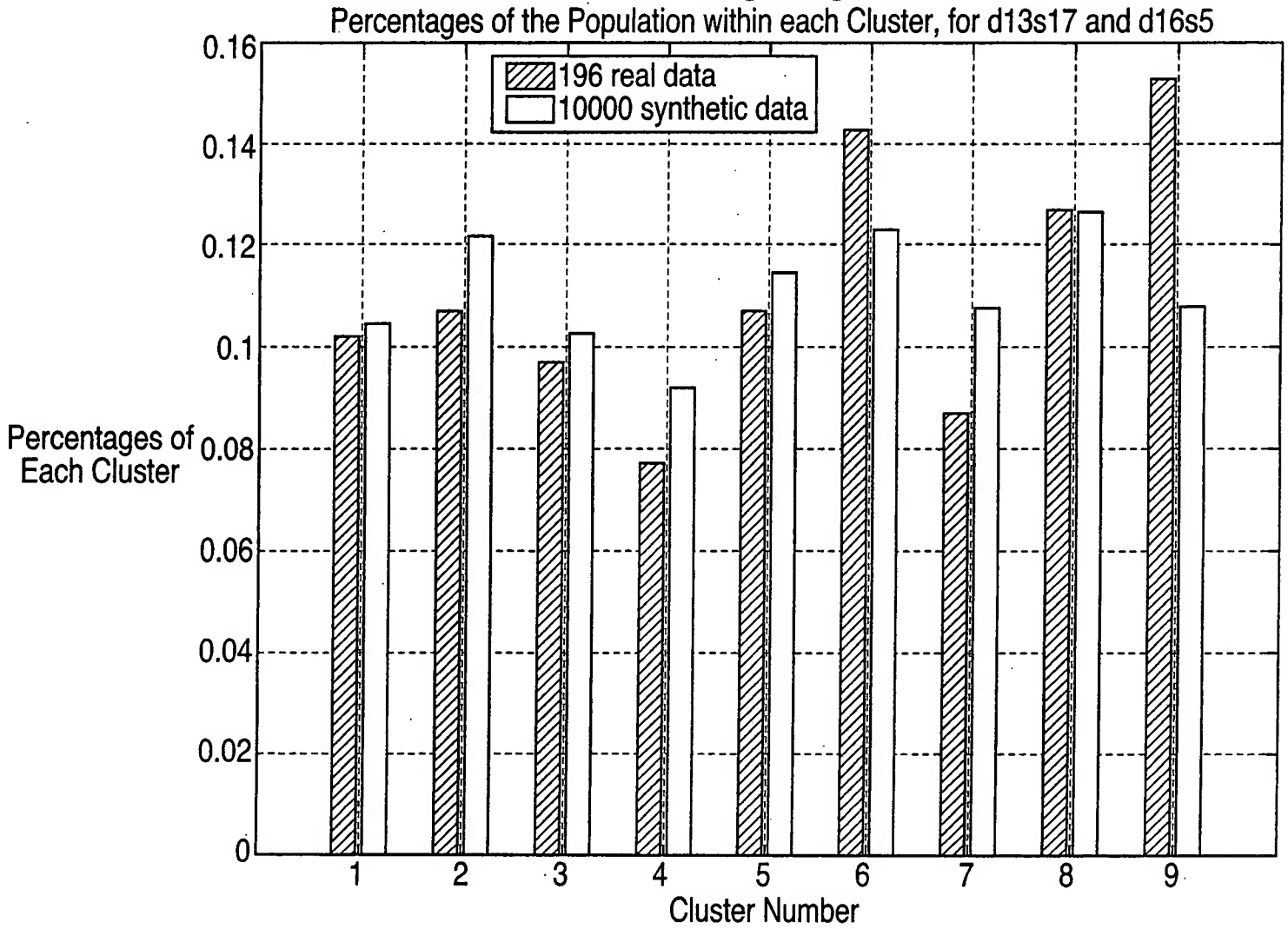

Percentages of the Population within each Cluster, for d13s17 and d16s5
196 real data
10000 synthetic data
Percentages of Each Cluster
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
1
2
3
4
5
6
7
8
9
Cluster Number

FIG. 26

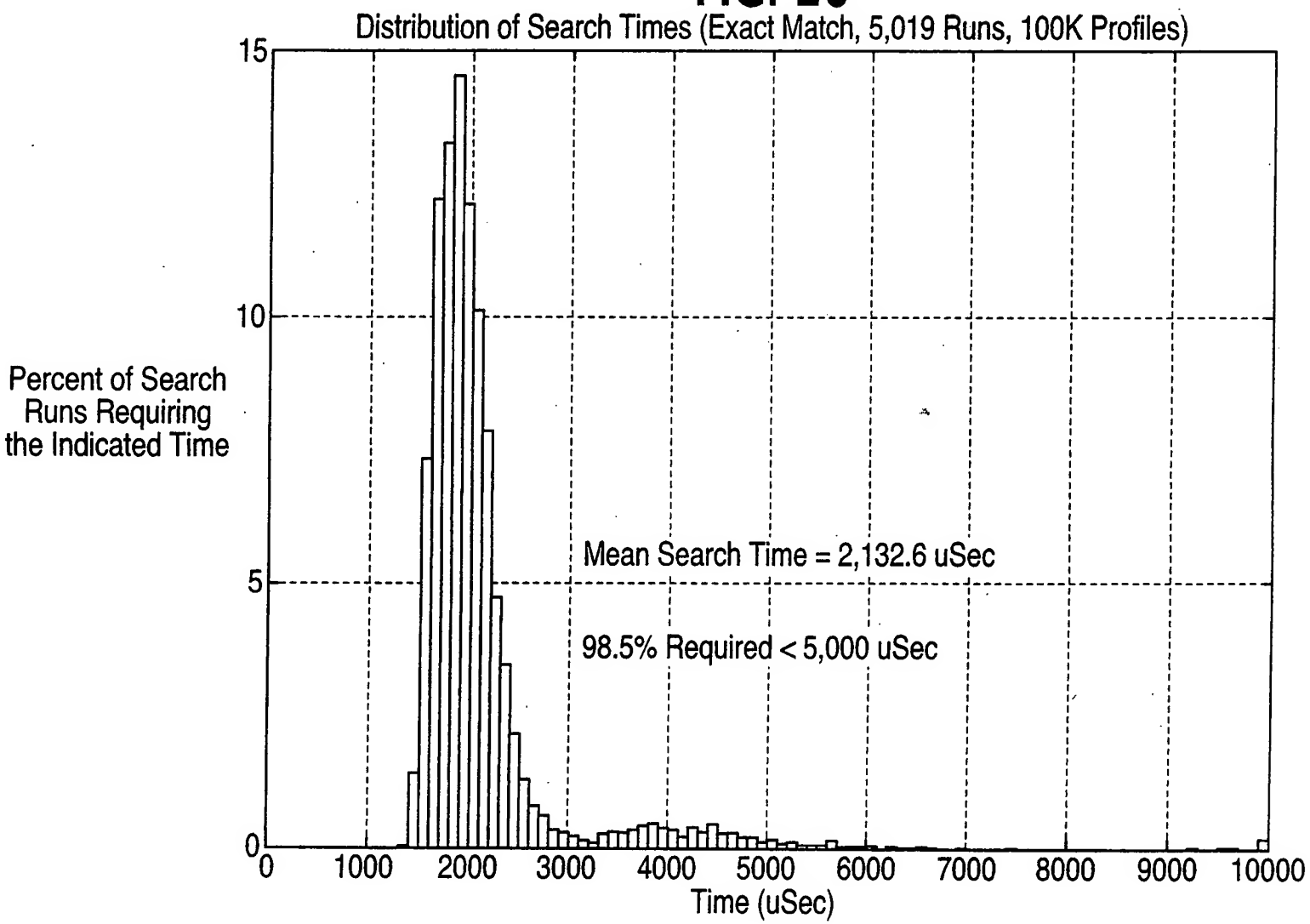

Distribution of Search Times (Exact Match, 5,019 Runs, 100K Profiles)
Percent of Search Runs Requiring the Indicated Time
15
10
5
0
Mean Search Time = 2,132.6 uSec
98.5% Required < 5,000 uSec
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
10000
Time (uSec)

FIG. 27

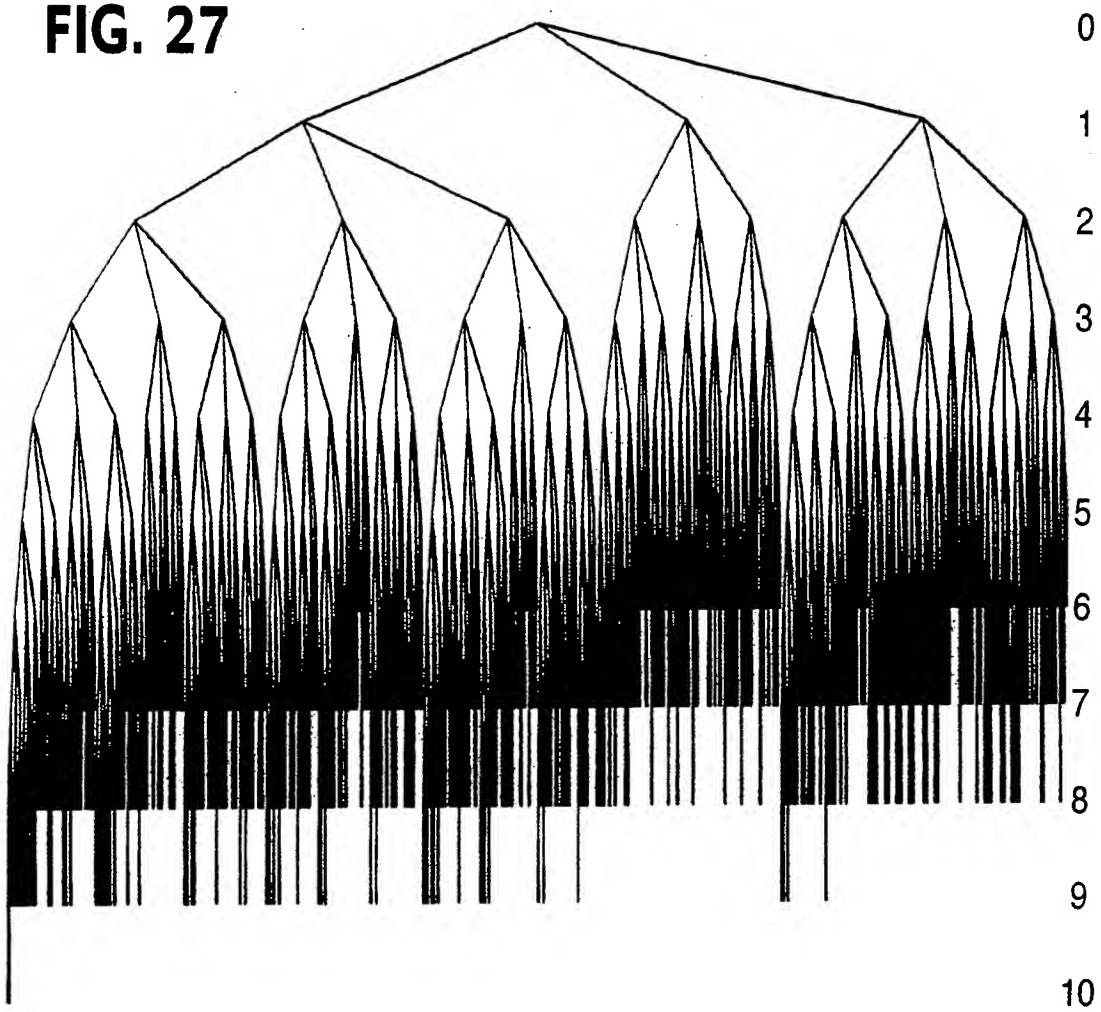
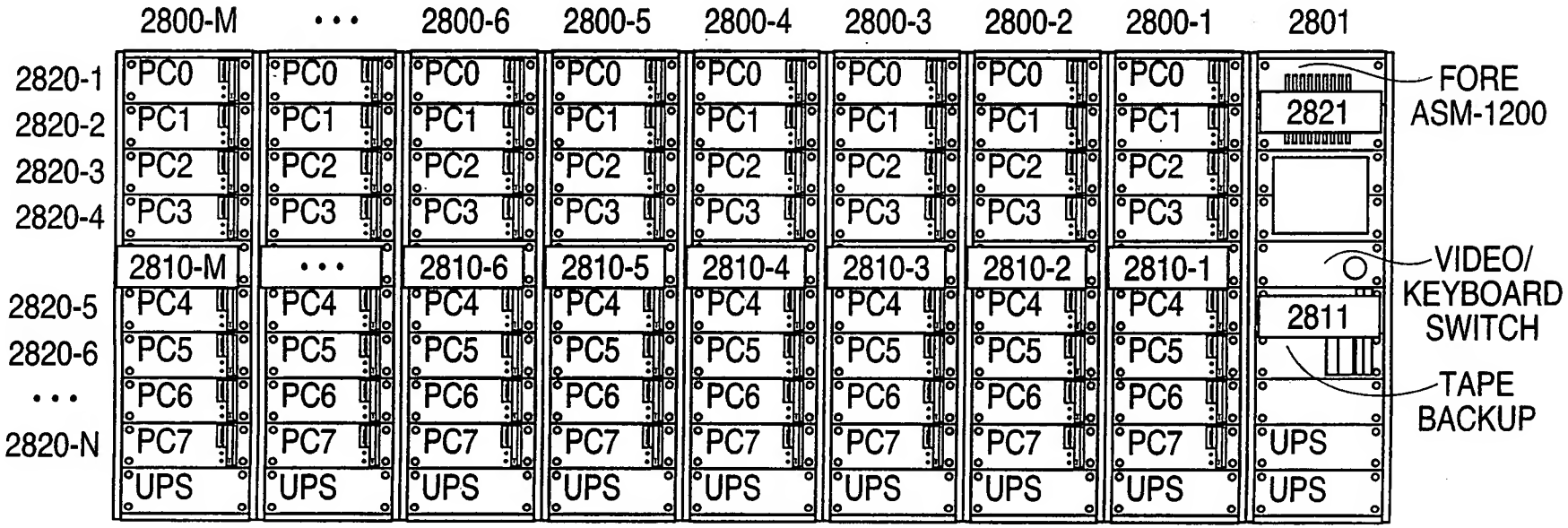


FIG. 28


2800-M
• • •
2800-6
2800-5
2800-4
2800-3
2800-2
2800-1
2801
2820-1
2820-2
2820-3
2820-4
2820-5
2820-6
• • •
2820-N
PC0
PC1
PC2
PC3
2810-M
• • •
2810-6
2810-5
2810-4
2810-3
2810-2
2810-1
PC4
PC5
PC6
PC7
UPS
2821
2811
FORE
ASM-1200
VIDEO/
KEYBOARD
SWITCH
TAPE
BACKUP